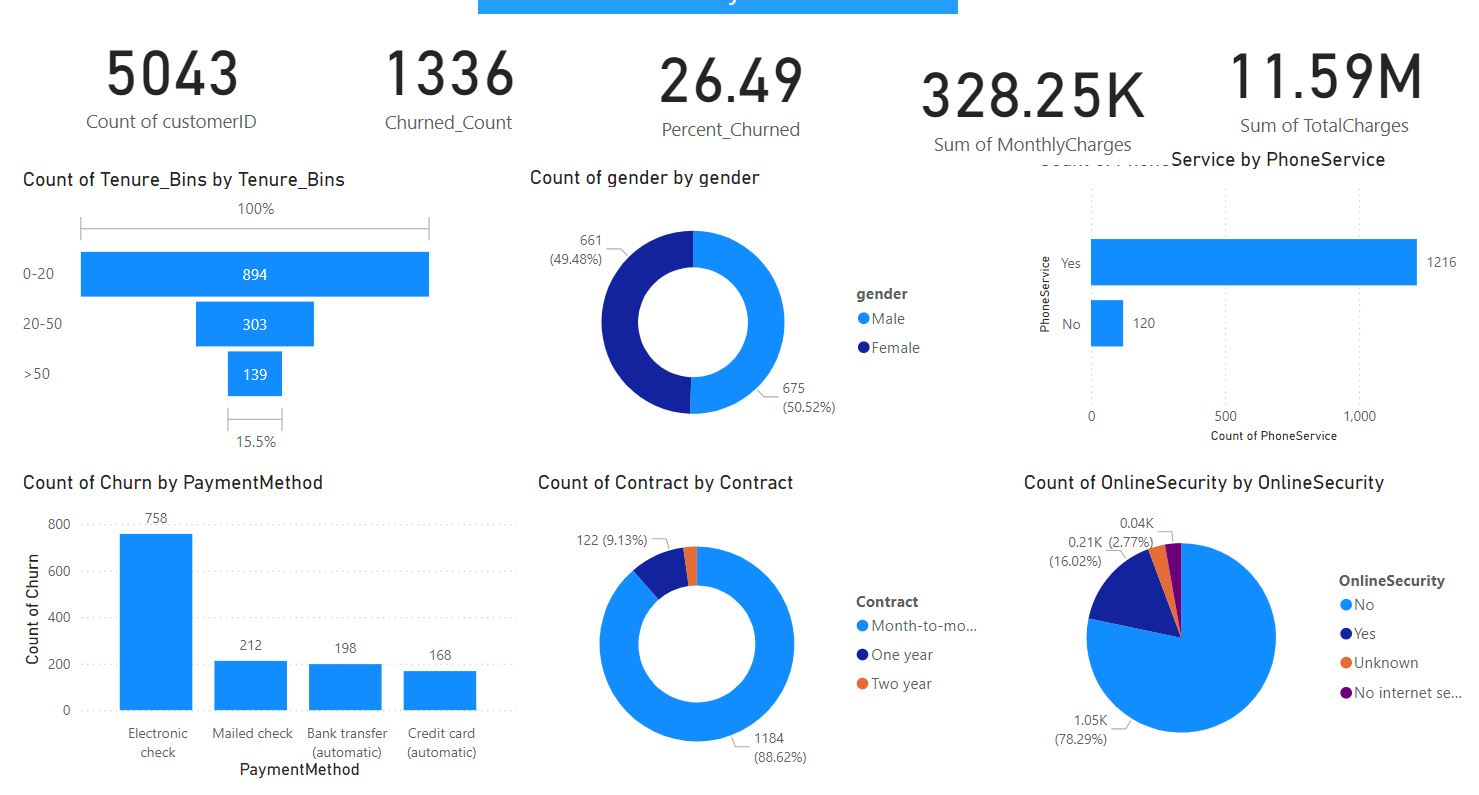
**Understanding Customer Churn in the Telecommunication Industry: A Comprehensive Analysis**

****

**Introduction**

In today's competitive market, comprehending the dynamics of customer churn is essential for businesses aiming to enhance retention and mitigate losses. One pivotal factor influencing churn rates is the type of contract a customer holds. In this blog post, we will delve into the statistical relationship between churn rates and contract terms using a Chi-Square test, followed by the development and implementation of machine learning models to predict which customers are likely to leave the network. These models will utilize features such as contract type, tenure, phone service, internet service, paperless billing, and the ratio of monthly charges to total charges, among others. Notably, these features have demonstrated high predictive importance in logistic regression models for churn prediction.

**Methodology**

**Project Background and Objective**

In the ever-evolving telecommunication sector, retaining customers and accurately forecasting churn is essential for sustaining growth and profitability. This project seeks to harness advanced classification models to deliver a thorough understanding of customer behaviour for a telecommunication company. The project focuses on:

Churn Prediction: Develop and implement machine learning models to predict which customers are likely to leave the network.

**Stakeholders and Key Metrics**

Stakeholders: Telecommunication provider

Success Criteria:

Model accuracy of 85% (on balanced data)

F1 score > 80%

ROC Curve > 80%

Minimum of 2 baseline models with hyperparameter tuning applied only if they exceed the scores above.

**Data and Features**

The dataset includes features such as customerID, gender, SeniorCitizen, Partner, Dependents, tenure, PhoneService, MultipleLines, InternetService, OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV, StreamingMovies, Contract, PaperlessBilling, PaymentMethod, MonthlyCharges, TotalCharges, and Churn.

**Hypothesis Testing**

Null Hypothesis (H0): There is no statistically significant relationship between churn rate and the contract term of the customer.

Alternative Hypothesis (H1): There is a statistically significant relationship between churn rate and the contract term of the customer.

**Models Used**

K Neighbors Classifier (KNN)

Logistic Regression Classifier (LR)

Random Forest Classifier (RF)

Support Vector Machine Classifier (SVM)

Gradient Boosting Classifier (GB)

XGBoost Classifier (XB)

**Deliverables**

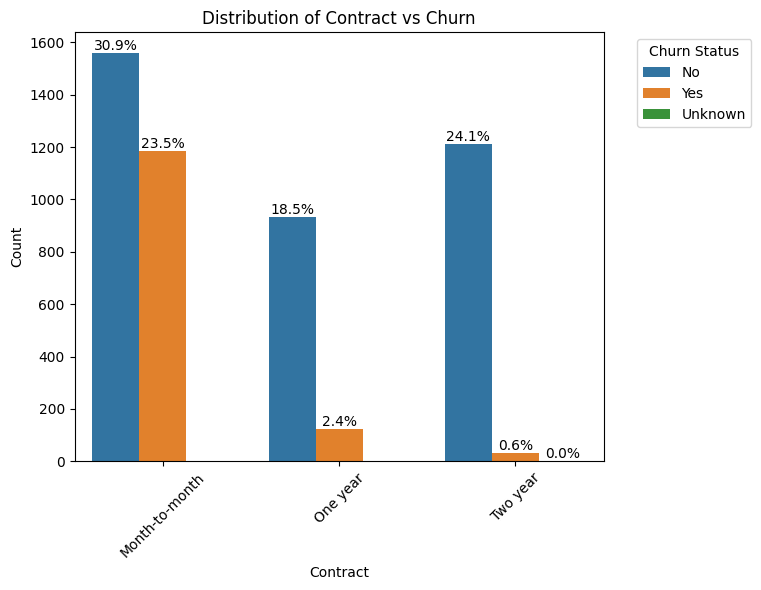
Power BI Dashboard: Visual representation of insights and findings.

GitHub Repository: Code and documentation for reproducibility and collaboration.

**Key Insights**

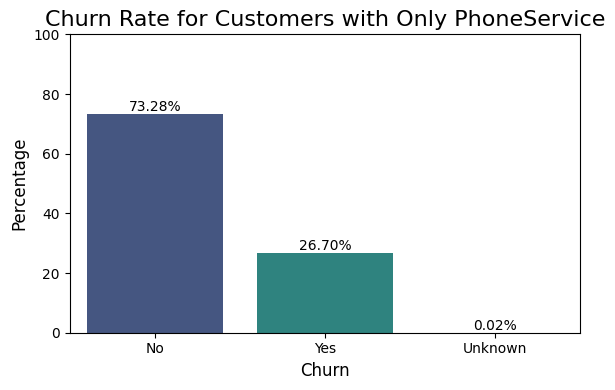
**Impact of Contract Type on Churn Rate**

Analysing how different contract types (month-to-month, one-year, two-year) affect churn rates can provide insights into customer loyalty and satisfaction. The two-year contract has the lowest churn rate compared to the one-year contract. However, the month-to-month contract, with the highest churn rate, is not too distinct from the company churn rate of 26.5%, indicating a higher propensity to churn among customers with more flexible contracts.



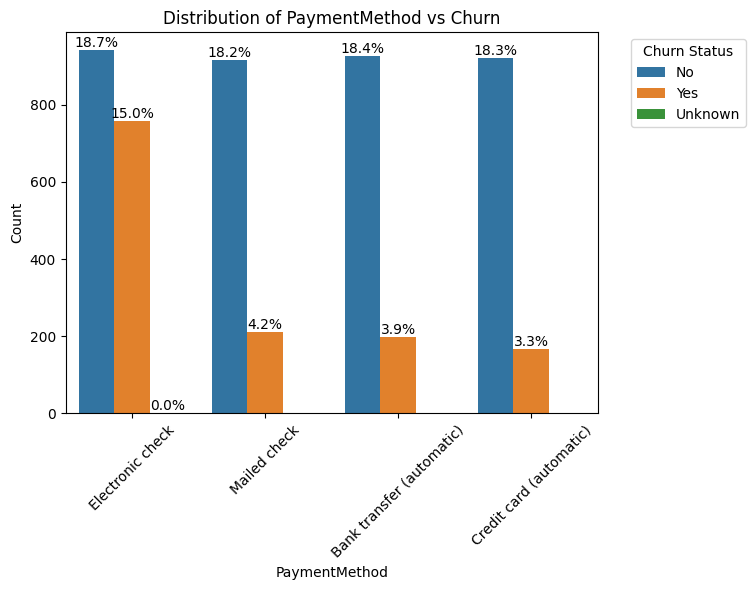
**Churn Rate for Customers with Only Phone Service**

Understanding the churn rate among customers with only phone service can help identify if this specific segment is more prone to leaving compared to customers with other or additional services. The churn rate for customers with only phone service is 26.7%, which is very close to the overall telco churn rate of 26.5%. This indicates that phone-only customers churn at a similar rate to the general customer base.



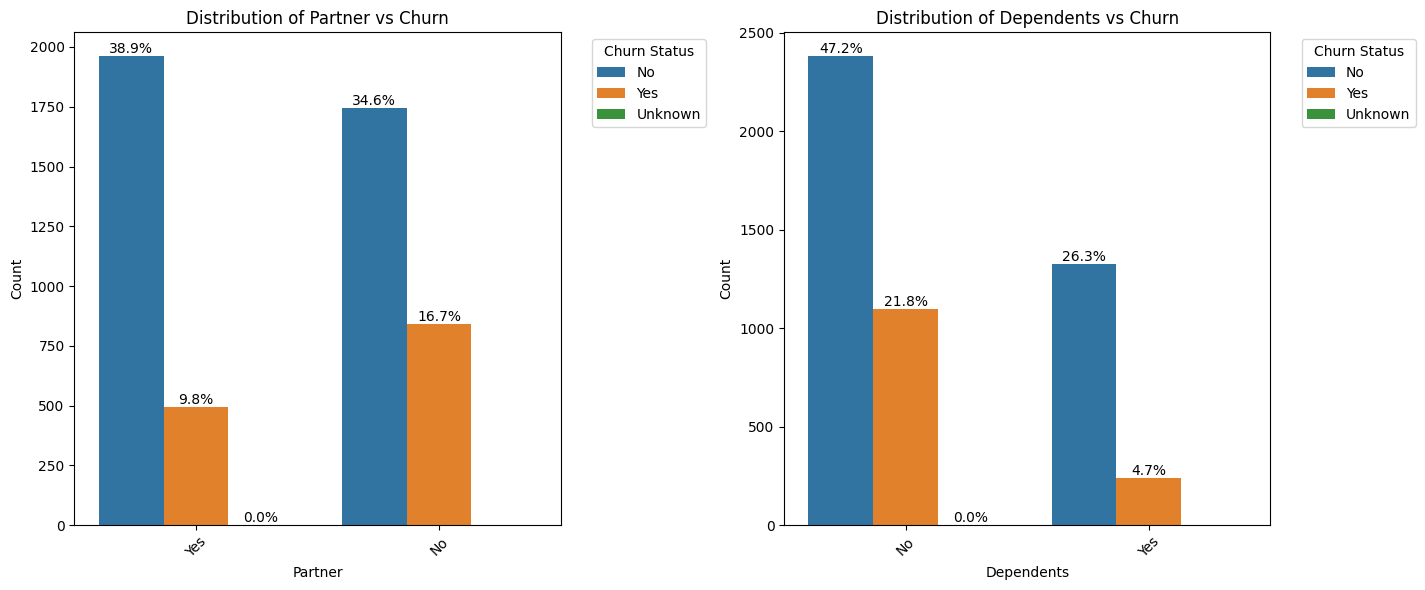
**Customer Payment Method and Churn Rate**

Certain payment methods might be more convenient or reliable for customers, potentially influencing their decision to stay or leave. The churn rate is highest for subscribers using electronic check payments, with a churn rate of 44%, followed by mailed checks and bank transfers. This suggests that electronic check payments may be less reliable or convenient, potentially leading to dissatisfaction and higher churn.



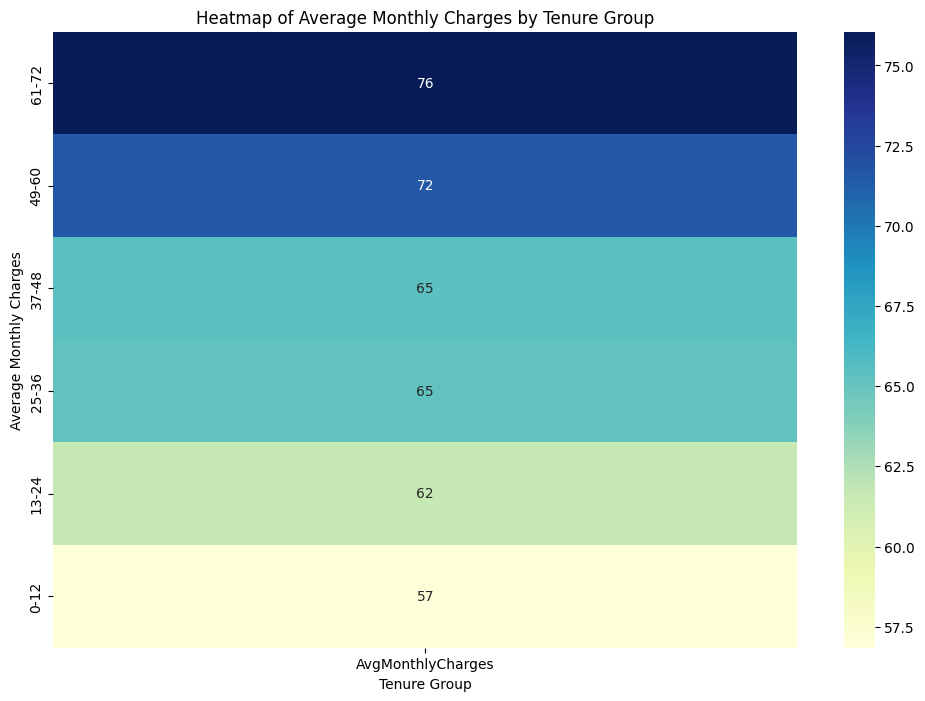
**Impact of Partner or Dependents on Churn**

The presence of a partner or dependents might influence a customer’s stability and satisfaction with the service, impacting churn. Subscribers without partners have a higher churn rate (16.7%) in contrast with those having partners (9.8%). Similarly, subscribers without dependents have a higher churn rate (21.8%) compared to those with dependents (4.7%). These insights indicate that the presence of a partner or dependents is associated with lower churn rates.



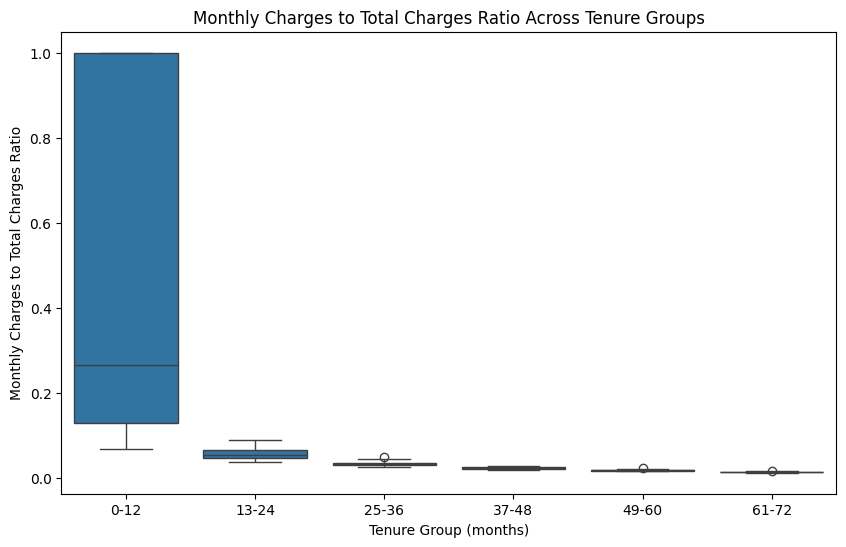
**Average Monthly Charges Across Different Customer Tenures**

Examining the variation in average monthly charges across different customer tenures can provide insights into spending behaviour over time. For instance, long-term customers might have different usage patterns and billing amounts compared to new customers. Feature engineering was performed to create the 'AvgMonthlyCharges' feature, defined as the total charges divided by tenure. Customers in the early tenure groups (0-12 and 13-24 months) are at higher risk of churn, possibly due to lower perceived value in the services for the cost they are incurring. Mid-tenure groups (25-36 and 37-48 months) show stability in charges, which may indicate a balanced relationship with the service provider, reducing churn risk. Long-tenure groups (49-60 and 61-72 months) with higher average monthly charges demonstrate strong customer loyalty and a lower risk of churn, as they likely perceive significant value in the services provided.



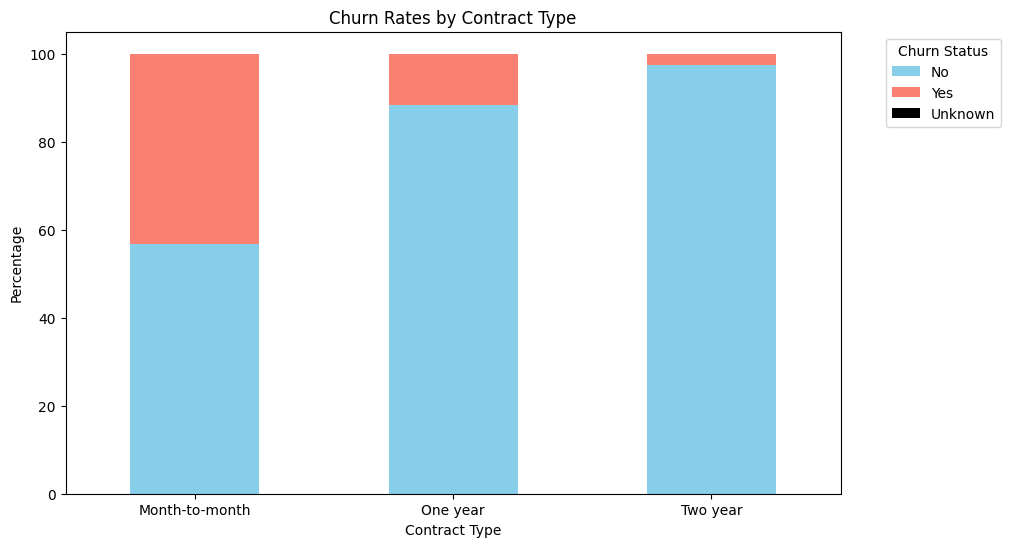
**Monthly Charges to Total Charges Ratio**

Analysing the ratio of monthly charges to total charges can reveal spending patterns and financial commitment over a customer’s tenure. Feature engineering was performed to create the 'MonthlyChargesToTotalChargesRatio' feature, defined as the monthly charges divided by total charges. Higher churn risk is observed among customers in the 0-12 months group, with high ratios and wide interquartile ranges (IQR), indicating they are new and might not have found sufficient value in the services yet. Moderate churn risk is seen in the 13-24 months group, showing more stability but still at risk if they do not perceive long-term value. Lower churn risk is evident in the 25-36, 37-48, 49-60-, and 61-72-months groups, with very low ratios and minimal variation, indicating strong commitment and satisfaction with the services, thereby showing the lowest likelihood of churn.



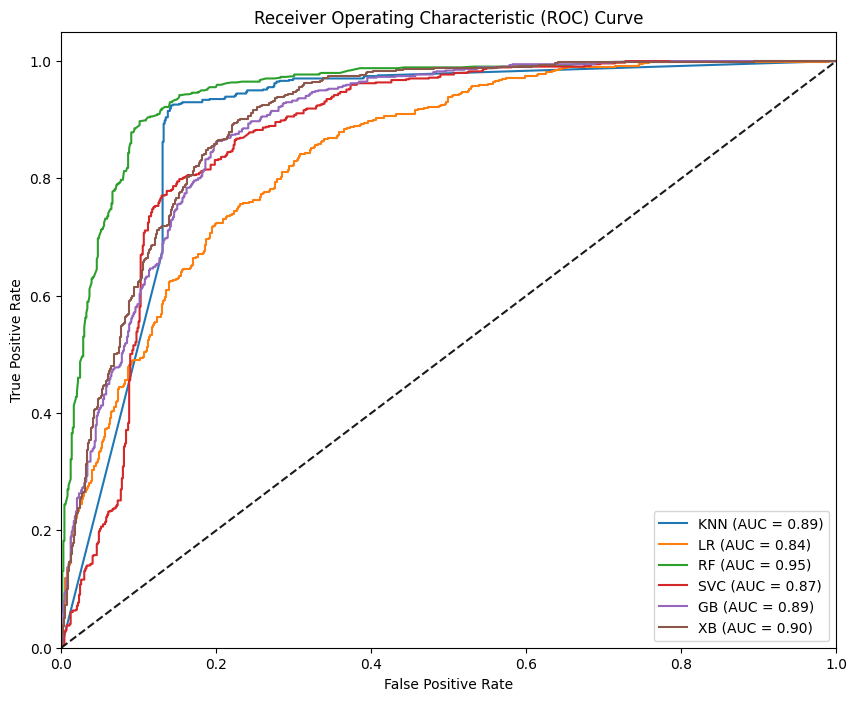
**Hypothesis Support**

The insights derived from the data support the hypothesis that there is a statistically significant relationship between churn rate and the contract term of the customer. The Chi-Square test results with an extremely low p-value provide strong evidence against the null hypothesis, confirming the significant relationship.



**Business Impact Assessment and Documentation of the Model**

**Model Performance**

****

Best Model: Random Forest (RF)

Reason: Offers the highest ROC AUC score, high recall, excellent F1 score, and low log loss.

Second Best Model: XGBoost (XB)

Reason: Consistently performs well across all metrics, making it a reliable alternative to Random Forest.

**Practical Implications**

High Recall: Ensures that most churners are identified, allowing for effective intervention and reduced churn rate.

Low Log Loss: Provides reliable probabilistic predictions, enabling better decision-making and resource allocation.

Balanced Metrics: Ensures interventions are both effective and efficient, minimizing unnecessary costs while maximizing retention.

**Documentation**

The project is meticulously documented for future reference and ease of implementation.

Deliverables include:

[Power BI Dashboard](https://app.powerbi.com/view?r=eyJrIjoiNzFlMTM2ZTAtNWUxZC00Y2EzLWJhZGQtNjUzZTg0NGYxOWY2IiwidCI6IjQ0ODdiNTJmLWYxMTgtNDgzMC1iNDlkLTNjMjk4Y2I3MTA3NSJ9)

[GitHub Repository](https://github.com/Nfayem/Telco_Churn_ML.git)

**Conclusion**

By using a Chi-Square test to analyze the relationship between customer churn and contract terms, we have uncovered a statistically significant association. This knowledge allows businesses to make data-driven decisions to improve customer retention. By focusing on contract types that are more likely to lead to churn, companies can develop more effective strategies to keep their customers satisfied and loyal.

Understanding the nuances of customer behavior is key to staying competitive, and leveraging statistical insights is a powerful way to achieve this. As we have seen, even seemingly straightforward data points like contract terms can provide valuable insights that drive business success.

**Appreciation**

I highly recommend Azubi Africa for their comprehensive and effective programs. Read More articles about Azubi Africa here and take a few minutes to visit this link (<https://www.azubiafrica.org/>) to learn more about Azubi Africa life changing programs.